Diffraction of Apodized Pupil Mapping Systems with Aberrations

Ruslan Belikov

Mechanical and Aerospace Engineering, Princeton University

rbelikov@princeton.edu

N. Jeremy Kasdin

Mechanical and Aerospace Engineering, Princeton University

jkasdin@princeton.edu

Robert J. Vanderbei

Operations Research and Financial Engineering, Princeton University

rvdb@princeton.edu

ABSTRACT

Pupil mapping is a promising and unconventional new method for high contrast imaging being considered for terrestrial exoplanet searches. It employs two (or more) specially designed aspheric mirrors to create a high-contrast amplitude profile across the telescope pupil that does not appreciably attenuate amplitude. As such, it reaps significant benefits in light collecting efficiency and inner working angle, both critical parameters for terrestrial planet detection. While much has been published on various aspects of pupil mapping systems, the problem of sensitivity to wavefront aberrations remains an open question. In this paper, we present an efficient method for computing the sensitivity of a pupil mapped system to Zernike aberrations. We then use this method to study the sensitivity of a particular pupil mapping system and compare it to the concentric-ring shaped pupil coronagraph. In particular, we show how contrast and inner working angle degrade with increasing Zernike order and rms amplitude, which has obvious ramifications for the stability requirements and overall design of a planet-finding observatory.

Subject headings: Extrasolar planets, coronagraphy, Fresnel propagation, diffraction analysis, point spread function, pupil mapping, apodization, PIAA

1. Introduction

The impressive discoveries of large extrasolar planets over the past decade has inspired widespread interest in finding and directly imaging Earth-like planets in the habitable zones of nearby stars. In fact, NASA has plans to launch two space telescopes to accomplish this, the Terrestrial Planet Finder Coronagraph (TPF-C) and the Terrestrial Planet Finder Interferometer (TPF-I), while the European Space Agency is planning a similar multi-satellite mission called Darwin. These missions are currently in the concept study phase. In addition, numerous ground-based searches are proceeding using both coronagraphic and interferometric approaches.

Direct imaging of Earth-like extrasolar planets poses extremely challenging problems. To see why, consider viewing our solar system from a good distance (say from another nearby star system), our Sun would appear 10^{10} times brighter than Earth. Hence, we need an imaging system capable of detecting planets that are 10 orders of magnitude fainter than the star they orbit. Such a system is referred to as a high-contrast imager. Furthermore, given the distances involved, the angular separation for most targets is very small, requiring the largest launchable telescope possible.

For TPF-C, for example, the current baseline design involves a traditional *Lyot coronagraph* consisting of a modern 8th-order occulting mask attached to the back end of a Ritchey-Chretien telescope having an 8m by 3.5m elliptical primary mirror (see, e.g., Kuchner et al. (2004)). Alternative innovative back-end designs still being considered include *shaped pupils* (see, e.g., Kasdin et al. (2003) and Vanderbei et al. (2004)), a *visible nuller* (see, e.g., Shao et al. (2004)) and *pupil mapping* (see, e.g., Guyon (2003) where this technique is called phase-induced amplitude apodization or PIAA). By pupil mapping we mean a system of two lenses, or mirrors, that take a flat input field at the entrance pupil and produce an output field that is amplitude modified but still flat in phase (at least for on-axis sources).

Pupil mapping has received considerable attention recently because of its high throughput and small effective inner working angle (IWA). These benefits could potentially permit more observations over the mission lifetime, or conversely, a smaller and cheaper overall telescope. As a result, there have been numerous studies over the past year to examine the performance of pupil mapping systems. In particular, in Traub and Vanderbei (2003) and Vanderbei and Traub (2005), formulas for the optical surfaces were derived using ray optics. However, this analysis failed to account for the complete diffraction through a pupil mapping system. More recently, Vanderbei (2006) provided a detailed diffraction analysis. Unfortunately, this analysis showed that a pure pupil mapping system cannot achieve the needed 10^{-10} contrast; the diffraction effects from the pupil mapping systems themselves are

so detrimental that contrast attained at the first image plane is limited to 10^{-5}. In Guyon et al. (2005) and Pluzhnik et al. (2006), a hybrid pupil mapping system was proposed that combines the pupil mapping mirrors with a slight apodization of oversized entrance and exit pupils. This combination does indeed achieve the needed high-contrast point spread function (PSF). In what follows, we will refer to such apodized pupil mapping simply as *pupil mapping*, and use the term *pure pupil mapping* to refer to the original unapodized system.

What remains to be answered is how pupil mapping behaves in the presence of optical aberrations. It is essential that contrast be maintained during an observation while the system is being aberrated due to the small dynamic perturbations of the primary mirror. An understanding of this sensitivity is critical to the design of TPF-C or any other observatory. In Green et al. (2004), a detailed sensitivity analysis of both shaped pupils and Lyot coronagraphs is given. Kuchner et al. (2004) introduced an 8th order image plane mask for the Lyot coronagraph that reduces the sensitivity to low-order aberrations to the level of shaped pupils. Both of these design approaches achieve the needed sensitivity for a realizable mission. Up to now, however, no comparable study has been done for pupil mapping.

In this paper, we present an efficient method for computing the effects of wavefront aberrations on pupil mapping. We begin with a brief review of the design of pupil mapping systems in Section 2. We then present in Section 3.1 a semi-analytical approach to computing the PSF of a pupil-mapping system in the presence of aberrations represented by Zernike polynomials. The assumption of Zernikes allows an analytical treatment of the azimuthal angle, reducing the computational problem from a double integral to a few single ones, thus eliminating the need for massive computing power. We also show how these methods can be used to compute PSFs for a purely apodizing coronagraph in Section 3.2 as well as off-axis PSFs in Section 3.3. We conclude Section 3 with a sensitivity analysis of pupil mapping systems in Section 3.4, showing the degradation of contrast and inner working angle with Zernike order and rms amplitude. We also demonstrate the achievable contrast limit of the current pupil mapping approach. In Section 4, we repeat the sensitivity analysis for an alternative coronagraph design, the shaped pupil coronagraph with concentric ring masks, and compare to pupil mapping. In particular, we show that the pupil mapping approach to high-contrast is more sensitive to aberrations than a purely apodized one, possibly making it more difficult to exploit the inner working angle advantage of pupil mapping.

2. Review of Pupil Mapping and Apodization

2.1. Pure Pupil Mapping via Ray Optics

We begin by summarizing the ray-optics description of pure pupil mapping. An on-axis ray entering the first pupil at radius r from the center is to be mapped to radius $\tilde{r} = \tilde{R}(r)$ at the exit pupil. Optical elements at the two pupils ensure that the exit ray is parallel to the entering ray. The function $\tilde{R}(r)$ is assumed to be positive and increasing or, sometimes, negative and decreasing. In either case, the function has an inverse that allows us to recapture r as a function of $\tilde{r}$: $r = R(\tilde{r})$. The purpose of pupil mapping is to create nontrivial amplitude profiles. An amplitude profile function $A(\tilde{r})$ specifies the ratio between the output amplitude at $\tilde{r}$ to the input amplitude at r (in a pure pupil-mapping system the input amplitude is constant). Vanderbei and Traub (2005) showed that for any amplitude profile $A(\tilde{r})$ there is a pupil mapping function $R(\tilde{r})$ that achieves this profile. Specifically, the pupil mapping is given by

$$R(\tilde{r}) = \pm\sqrt{\int_0^{\tilde{r}} 2A^2(s)s\,ds}. \tag{1}$$

Furthermore, if we consider the case of a pair of lenses that are plano on their outward-facing surfaces, then the inward-facing surface profiles, $h(r)$ and $\tilde{h}(\tilde{r})$, that are required to obtain the desired pupil mapping are given by the solutions to the following ordinary differential equations:

$$\frac{\partial h}{\partial r}(r) = \frac{r - \tilde{R}(r)}{\sqrt{Q_0^2 + (n^2 - 1)(r - \tilde{R}(r))^2}}, \qquad h(0) = z, \tag{2}$$

and

$$\frac{\partial \tilde{h}}{\partial \tilde{r}}(\tilde{r}) = \frac{R(\tilde{r}) - \tilde{r}}{\sqrt{Q_0^2 + (n^2 - 1)(R(\tilde{r}) - \tilde{r})^2}}, \qquad \tilde{h}(0) = 0. \tag{3}$$

Here, n is the refractive index and Q_0 is a constant determined by the distance z separating the centers ($r = 0$, $\tilde{r} = 0$) of the two lenses: $Q_0 = -(n-1)z$.

Let $S(r, \tilde{r})$ denote the distance between a point on the first lens surface r units from the center and the corresponding point on the second lens surface $\tilde{r}$ units from its center. Up to an additive constant, the optical path length of a ray that exits at radius $\tilde{r}$ after entering at radius $r = R(\tilde{r})$ is given by

$$Q_0(\tilde{r}) = S(R(\tilde{r}), \tilde{r}) + n(\tilde{h}(\tilde{r}) - h(R(\tilde{r}))). \tag{4}$$

Vanderbei and Traub (2005) showed that, for an on-axis source, $Q_0(\tilde{r})$ is constant and equal to Q_0.

2.2. High-Contrast Amplitude Profiles

If we assume that a collimated beam with amplitude profile $A(\tilde{r})$ such as one obtains as the output of a pupil mapping system is passed into an ideal imaging system with focal length f, the electric field $E(\rho)$ at the image plane is given by the Fourier transform of $A(\tilde{r})$:

$$E(\xi, \eta) = \frac{E_0}{\lambda f} \int \int e^{-2\pi i \frac{\tilde{x}\xi + \tilde{y}\eta}{\lambda f}} A(\sqrt{\tilde{x}^2 + \tilde{y}^2}) d\tilde{y} d\tilde{x}. \tag{5}$$

Here, E_0 is the input amplitude which, unless otherwise noted, we take to be unity. Since the optics are azimuthally symmetric, it is convenient to use polar coordinates. The amplitude profile A is a function of $\tilde{r} = \sqrt{\tilde{x}^2 + \tilde{y}^2}$ and the image-plane electric field depends only on image-plane radius $\rho = \sqrt{\xi^2 + \eta^2}$:

$$E(\rho) = \frac{1}{\lambda f} \int \int e^{-2\pi i \frac{\tilde{r}\rho}{\lambda f} \cos(\theta - \phi)} A(\tilde{r}) \tilde{r} d\theta d\tilde{r} \tag{6}$$

$$= \frac{2\pi}{\lambda f} \int J_0 \left(-2\pi \frac{\tilde{r}\rho}{\lambda f} \right) A(\tilde{r}) \tilde{r} d\tilde{r}. \tag{7}$$

The point-spread function (PSF) is the square of the electric field:

$$\text{Psf}(\rho) = |E(\rho)|^2. \tag{8}$$

For the purpose of terrestrial planet finding, it is important to construct an amplitude profile for which the PSF at small nonzero angles is ten orders of magnitude reduced from its value at zero. Vanderbei et al. (2003a) explains how these functions are computed as solutions to certain optimization problems.

The high-contrast amplitude profile used in the rest of this paper is shown in Figure 1.

We end this section by noting that the need to design customized amplitude profiles arise in many areas, usually in the context of apodization—i.e., profiles that only attenuate the beam. Slepian (1965) was perhaps the first to study this problem carefully. For some recent applications, the reader is refered to the following papers in the area of beam shaping: Carney and Gbur (1999), Goncharov et al. (2002) and Hoffnagle and Jefferson (2003).

2.3. Apodized Pupil Mapping Systems

Vanderbei (2006) showed that pure pupil mapping systems designed for contrast of 10^{-10} actually achieve much less than this due to harmful diffraction effects that are not captured by the simple ray tracing analysis outlined in the previous section. For most systems of

practical real-world interest (i.e., systems with apertures of a few inches and designed for visible light), contrast is limited to about 10^{-5}. In Vanderbei (2006), certain hybrid designs were considered that improve on this level of performance but none of the hybrid designs presented there completely overcame this diffraction-induced contrast degradation.

In this section, we describe an apodized pupil mapping system that is somewhat more complicated than the designs presented in Vanderbei (2006). This hybrid design, based on ideas proposed by Olivier Guyon and Eugene Pluzhnik (see Pluzhnik et al. (2006)), involves three additional components. They are

1. a preapodizer A_0 to soften the edge of the first lens/mirror so as to minimize diffraction effects caused by hard edges,

2. a postapodizer to smooth out low spatial frequency ripples produced by diffraction effects induced by the pupil mapping system itself, and

3. a backend phase shifter to smooth out low spatial frequency ripples in phase.

Note that the backend phase shifter can be built into the second lens/mirror. There are several choices for the preapodizer. For this paper, we use the preapodizer given by Eqs. (3) and (4) in Pluzhnik et al. (2006):

$$A_0(r) = \frac{A(r)(1+\beta)}{A(r) + \beta A_{\max}},$$

where $A_{\max}$ denotes the maximum value of $A(r)$ and β is a scalar parameter, which we take to be 0.1. It is easy to see that

- $A(r)/A_{\max} \leq A_0(r) \leq 1$ for all r,

- $A_0(r)$ approaches 1 as $A(r)$ approaches $A_{\max}$, and

- $A_0(r)$ approaches 0 as $A(r)$ approaches 0.

Incorporating a post-apodizer introduces a degree of freedom that is lacking in a pure pupil mapping system. Namely, it is possible to design the pupil mapping system based on an arbitrary amplitude profile and then convert this profile to a high-contrast profile via an appropriate choice of backend apodizer. We have found that a simple Gaussian amplitude profile that approximately matches a high-contrast profile works very well. Specifically, we used

$$A_{\text{pupmap}}(\tilde{r}) = 3.35 e^{-22(\tilde{r}/\tilde{a})^2},$$

where $\tilde{a}$ denotes the radius of the second lens/mirror.

The backend apodization is computed by taking the actual output amplitude profile as computed by a careful diffraction analysis, smoothing it by convolution with a Gaussian distribution, and then apodizing according to the ratio of the desired high-contrast amplitude profile $A(\tilde{r})$ divided by the smoothed output profile. Of course, since a true apodization can never intensify a beam, this ratio must be further scaled down so that it is nowhere greater than unity. The Gaussian convolution kernel we used has mean zero and standard deviation $\tilde{a}/\sqrt{100,000}$.

The backend phase modification is computed by a similar smoothing operation applied to the output phase profile. Of course, the smoothed output phase profile (measured in radians) must be converted to a surface profile (having units of length). This conversion requires us to assume a certain specific wavelength. As a consequence, the resulting design is correct only at one wavelength. The ability of the system to achieve high contrast degrades as one moves away from the design wavelength.

Figure 2 shows plots characterizing the performance of the apodized pupil mapping system described in this section. The specifications for this system are as follows. The designed-for wavelength is 632.8nm. The optical elements are assumed to be mirrors separated by 0.375m. The system is an on-axis system and we therefore make the non-physical assumption that the mirrors don't obstruct the beam. That is, the mirrors are invisible except when they are needed. The mirrors take as input a 0.025m on-axis beam and produce a 0.025m pupil-remapped exit beam. The second mirror is oversized by a factor of two; that is, its diameter is 0.050m. The postapodizer ensures that only the central half contributes to the exit beam. The first mirror is also oversized appropriately as shown in the upper-right subplot of Figure 2.

After the second mirror, the exit beam is brought to a focus. The focal length is 2.5m. The lower-right subplot in Figure 2 shows the ideal PSF (in black) together with the achieved PSF at three wavelengths: at 70% (green), 100% (blue), and 130% (red) of the design wavelength. At the design wavelength, the achieved PSF matches the ideal PSF almost exactly. Note that there is minor degradation at the other two wavelengths mostly at low spatial frequencies.

It is important to note that the PSFs in Figure 2 correspond to a bright on-axis source (i.e., a star). Off-axis sources, such as faint planets, undergo two effects in a pupil mapping system that differ from the response of a conventional imaging system: an effective magnification and a distortion. These are explained in detail in Vanderbei and Traub (2005) and Traub and Vanderbei (2003). The magnification, in particular, is due to an effective

narrowing of the exit pupil as compared to the entrance pupil. It is this magnification that provides pupil mapped systems their smaller effective inner working angle. The techniques in Section 3 will allow us to compute the exact off-axis diffraction pattern of an apodized pupil mapped coronagraph and thus to see these effects.

While the effective magnification of a pupil mapping system results in an inner working angle advantage of about a factor of two, it does not produce high-quaity diffraction limited images of off-axis sources because of the distortion inherent in the system. Guyon (2003) proposed the following solution to this problem. He suggested using this system merely as a mechanism for concentrating (on-axis) starlight in an image plane. He then proposed that an occulter be placed in the image plane to remove the starlight. All other light, such as the distorted off-axis planet light, would be allowed to pass through the image plane. On the back side would be a second, identical pupil mapping system (with the apodizers removed), that would "umap" the off-axis beam and thus remove the distortions introduced by the first system (except for some beam walk—see Vanderbei and Traub (2005)).

3. Diffraction Analysis of Apodized Pupil Mapping with an Aberrated Wavefront

A general 2D diffraction analysis of a pupil mapping system with a possibly aberrated input field requires significant computing power. In this section, we present a much more efficient method assuming the aberrations are Zernike polynomials, which allows us to treat the azimuthal variables analytically. In particular, we perform a full end-to-end diffraction analysis of an apodized pupil mapping system with aberrations, as well as with a tilted field such as one from an off-axis planet. Such an analysis is required to determine the sensitivity of these systems to phase errors at the input pupil, as well as the off-axis response. At the end of this section, we describe the results for specific cases of interest.

3.1. Mathematical Development

If we assume that a flat, on-axis, electric field arrives at the entrance pupil, then the electric field at a particular point of the exit pupil can be well-approximated by superimposing the phase-shifted waves from each point across the entrance pupil (this is the well-known Huygens-Fresnel principle—see, e.g., Section 8.2 in Born and Wolf (1999)). For an apodized pupil mapping system, we can write this as

$$E_{\text{out}}(\tilde{x}, \tilde{y}) = A_{\text{out}}(\tilde{r}) \int \int \frac{1}{\lambda Q(\tilde{x}, \tilde{y}, x, y)} e^{2\pi i Q(\tilde{x}, \tilde{y}, x, y)/\lambda} A_{\text{in}}(r) dy dx, \qquad (9)$$

where

$$Q(\tilde{x}, \tilde{y}, x, y) = \sqrt{(x - \tilde{x})^2 + (y - \tilde{y})^2 + (h(r) - \tilde{h}(\tilde{r}))^2} + n(Z - h(r) + \tilde{h}(\tilde{r})) \qquad (10)$$

is the optical path length, Z is the distance between the plano lens surfaces (i.e., a constant slightly larger than z), $A_{\mathrm{in}}(r)$ denotes the input amplitude apodization at radius r, $A_{\mathrm{out}}(\tilde{r})$ denotes the output amplitude apodization at radius $\tilde{r}$, and where, of course, we have used r and $\tilde{r}$ as shorthands for the radii in the entrance and exit pupils, respectively. If the arriving field is not flat but instead has a phase profile given by $\alpha(x, y)$, then the integral must include this phase shift:

$$E_{\mathrm{out}}(\tilde{x}, \tilde{y}) = A_{\mathrm{out}}(\tilde{r}) \int \int \frac{1}{\lambda Q(\tilde{x}, \tilde{y}, x, y)} e^{2\pi i Q(\tilde{x}, \tilde{y}, x, y)/\lambda + i\alpha(x, y)} A_{\mathrm{in}}(r) dy dx. \qquad (11)$$

As before, it is convenient to work in polar coordinates:

$$E_{\mathrm{out}}(\tilde{r}, \tilde{\theta}) = A_{\mathrm{out}}(\tilde{r}) \int \int \frac{1}{\lambda Q(\tilde{r}, r, \theta - \tilde{\theta})} e^{2\pi i Q(\tilde{r}, r, \theta - \tilde{\theta})/\lambda + i\alpha(r, \theta)} A_{\mathrm{in}}(r) r d\theta dr, \qquad (12)$$

where

$$Q(\tilde{r}, r, \theta) = \sqrt{r^2 - 2r\tilde{r}\cos\theta + \tilde{r}^2 + (h(r) - \tilde{h}(\tilde{r}))^2} + n(Z - h(r) + \tilde{h}(\tilde{r})). \qquad (13)$$

For numerical tractability, it is essential to make approximations so that the integral over θ can be carried out analytically. To this end, we need to make an appropriate approximation to the square root term:

$$S = \sqrt{r^2 - 2r\tilde{r}\cos\theta + \tilde{r}^2 + (h(r) - \tilde{h}(\tilde{r}))^2}. \qquad (14)$$

We approximate the $1/Q(\tilde{r}, r, \theta - \tilde{\theta})$ amplitude-reduction factor in Eq. (12) by the constant $1/Z$ (the *paraxial approximation*). The $Q(\tilde{r}, r, \theta - \tilde{\theta})$ appearing in the exponential must, on the other hand, be treated with care. Recall that $Q(\tilde{r}, R(\tilde{r}), 0)$ is a constant. Since constant phase shifts are immaterial, we can subtract it from $Q(\tilde{r}, r, \theta)$ in Eq. (12) to get

$$E_{\mathrm{out}}(\tilde{r}, \tilde{\theta}) \approx \frac{A_{\mathrm{out}}(\tilde{r})}{\lambda Z} \int \int e^{2\pi i \left(Q(\tilde{r}, r, \theta - \tilde{\theta}) - Q(\tilde{r}, R(\tilde{r}), 0)\right)/\lambda + i\alpha(r, \theta)} A_{\mathrm{in}}(r) r d\theta dr. \qquad (15)$$

Next, we write the difference in Q's as follows:

$$
\begin{aligned}
Q(\tilde{r}, r, \theta - \tilde{\theta}) - Q(\tilde{r}, R(\tilde{r}), 0) &= S(\tilde{r}, r, \theta - \tilde{\theta}) - S(\tilde{r}, R(\tilde{r}), 0) + n(h(R(\tilde{r})) - h(r)) \\
&= \frac{S^2(\tilde{r}, r, \theta - \tilde{\theta}) - S^2(\tilde{r}, R(\tilde{r}), 0)}{S(\tilde{r}, r, \theta - \tilde{\theta}) + S(\tilde{r}, R(\tilde{r}), 0)} + n(h(R(\tilde{r})) - h(r))
\end{aligned}
\qquad (16)
$$

and then we expand out the numerator and cancel big terms that can be subtracted one from another to get

$$
\begin{aligned}
S^2(\tilde{r}, r, \theta - \tilde{\theta}) - S^2(\tilde{r}, R(\tilde{r}), 0) &= (r - R(\tilde{r}))(r + R(\tilde{r})) - 2\tilde{r}\left(r\cos(\theta - \tilde{\theta}) - R(\tilde{r})\right) \\
&\quad + (h(r) - h(R(\tilde{r})))\left(h(r) + h(R(\tilde{r})) - 2\tilde{h}(\tilde{r})\right). (17)
\end{aligned}
$$

When $r = R(\tilde{r})$ and $\theta = 0$, the right-hand side clearly vanishes as it should. Furthermore, for r close to $R(\tilde{r})$ and θ close to $\tilde{\theta}$, the right-hand side gives an accurate formula for computing the deviation from zero. That is to say, the right-hand side is easy to program in such a manner as to avoid subtracting one large number from another, which is always the biggest danger in numerical computation.

So far, everything is exact (except for the paraxial approximation). The only further approximation we make is to replace $S(\tilde{r}, r, \theta - \tilde{\theta})$ in the denominator of Eq. (16) with $S(\tilde{r}, R(\tilde{r}), 0)$ so that the denominator becomes just $2S(\tilde{r}, R(\tilde{r}), 0)$. Since $S(\tilde{r}, R(\tilde{r}), 0)$ appears many times in coming formulas, we abbreviate it as $S(\tilde{r})$ and hope this does not create any confusion. Putting this altogether, we get a new approximation, which we refer to as the *Huygens* approximation:

$$
\begin{aligned}
E_{\mathrm{out}}(\tilde{r}, \tilde{\theta}) &\approx \frac{A_{\mathrm{out}}(\tilde{r})}{\lambda Z} \int\int e^{2\pi i\left(\frac{(r-R(\tilde{r}))(r+R(\tilde{r}))+2\tilde{r}R(\tilde{r})+(h(r)-h(R(\tilde{r})))\left(h(r)+h(R(\tilde{r}))-2\tilde{h}(\tilde{r})\right)}{2S(\tilde{r})}+n(h(R(\tilde{r}))-h(r))\right)/\lambda} \\
&\quad \times e^{2\pi i\left(-\frac{\tilde{r}r\cos(\theta-\tilde{\theta})}{S(\tilde{r})}\right)/\lambda+i\alpha(r,\theta)} A_{\mathrm{in}}(r)\,d\theta\,r\,dr \qquad (18) \\
&= \frac{A_{\mathrm{out}}(\tilde{r})}{\lambda Z} \int K(r, \tilde{r}) L(r, \tilde{r}, \tilde{\theta}) A_{\mathrm{in}}(r)\,r\,dr, \qquad (19)
\end{aligned}
$$

where

$$
\begin{aligned}
K(r, \tilde{r}) &= e^{2\pi i\left(\frac{(r-R(\tilde{r}))(r+R(\tilde{r}))+2\tilde{r}R(\tilde{r})+(h(r)-h(R(\tilde{r})))\left(h(r)+h(R(\tilde{r}))-2\tilde{h}(\tilde{r})\right)}{2S(\tilde{r})}+n(h(R(\tilde{r}))-h(r))\right)/\lambda} \\
L(r, \tilde{r}, \tilde{\theta}) &= \int_0^{2\pi} e^{2\pi i\left(-\frac{\tilde{r}r\cos(\theta-\tilde{\theta})}{S(\tilde{r})}\right)/\lambda+i\alpha(r,\tilde{\theta})}\,d\theta. \qquad (20)
\end{aligned}
$$

Suppose now that the aberration is given by the (l, m)-th Zernike polynomial

$$
\alpha(r, \theta) = \epsilon Z_l^m(r/a)\cos(m\theta), \qquad (21)
$$

where ϵ is a small number. In this case, the kernel $L(r, \tilde{r}, \tilde{\theta})$ can be expressed in terms of Bessel functions. Recall that the n-th Bessel function $J_n(x)$ is defined by the requirement that $i^n J_n(x)$ be the n-th Fourier coefficient in the Fourier series expansion of $e^{ix\cos(\theta)}$. That is,

$$
J_n(x) = \frac{1}{2\pi i^n}\int_0^{2\pi} e^{ix\cos\theta} e^{in\theta}\,d\theta. \qquad (22)
$$

The Fourier series then, of course, is simply

$$e^{ix\cos\theta} = \sum_{k=-\infty}^{\infty} i^k J_k(x) e^{ik\theta}. \tag{23}$$

This series is usually referred to as the *Jacobi-Anger expansion*.

Plugging the Zernike polynomial into the definition of the kernel $L(r,\tilde{r},\tilde{\theta})$ and substituting a Jacobi-Anger expansion for each of the two cosine-exponentials, we get that

$$\begin{aligned}
L(r,\tilde{r},\tilde{\theta}) &= \int_0^{2\pi} e^{2\pi i\left(-\frac{\tilde{r}r\cos(\theta-\tilde{\theta})}{S(\tilde{r})}\right)/\lambda+i\alpha(r,\theta)} d\theta \tag{24}\\
&= \sum_j\sum_k \int i^j i^k J_j\left(-\frac{2\pi\tilde{r}r}{\lambda S(\tilde{r})}\right) e^{ij(\theta-\tilde{\theta})} J_k\left(\epsilon Z_l^m(r/a)\right) e^{ikm\theta} d\theta \\
&= \sum_j\sum_k i^{j+k} J_j\left(-\frac{2\pi\tilde{r}r}{\lambda S(\tilde{r})}\right) J_k\left(\epsilon Z_l^m(r/a)\right) e^{-ij\tilde{\theta}} \int e^{ij\theta+ikm\theta} d\theta \\
&= 2\pi \sum_k i^{k(1-m)} J_{km}\left(\frac{2\pi\tilde{r}r}{\lambda S(\tilde{r})}\right) J_k\left(\epsilon Z_l^m(r/a)\right) e^{ikm\tilde{\theta}} \\
&= \sum_k e^{ikm\tilde{\theta}} L_k(r,\tilde{r}), \tag{25}
\end{aligned}$$

where

$$L_k(r,\tilde{r}) = 2\pi i^{k(1-m)} J_{km}\left(\frac{2\pi\tilde{r}r}{\lambda S(\tilde{r})}\right) J_k\left(\epsilon Z_l^m(r/a)\right) \tag{26}$$

The second to the last equality above follows from the trivial fact that

$$\int_0^{2\pi} e^{in\theta} d\theta = \begin{cases} 2\pi & n=0 \\ 0 & \text{otherwise.} \end{cases}$$

Substituting Eq. (25) into Eq. (19), we can write

$$\begin{aligned}
E_{\text{out}}(\tilde{r},\tilde{\theta}) &= \frac{A_{\text{out}}(\tilde{r})}{\lambda Z} \sum_k e^{ikm\tilde{\theta}} \int K(r,\tilde{r}) L_k(r,\tilde{r}) A_{\text{in}}(r) r\, dr \\
&= \sum_k e^{ikm\tilde{\theta}} E_{\text{out},k}(\tilde{r}), \tag{27}
\end{aligned}$$

where

$$E_{\text{out},k}(\tilde{r}) = \frac{A_{\text{out}}(\tilde{r})}{\lambda Z} \int K(r,\tilde{r}) L_k(r,\tilde{r}) A_{\text{in}}(r) r\, dr. \tag{28}$$

Ignoring a complex unit exponential, the electric field in the image plane is just the Fourier transform of the electric field in the pupil plane. Working in polar coordinates, we

have

$$
\begin{aligned}
E_{\text{img}}(\rho, \phi) &= \frac{1}{\lambda f} \int \int e^{\frac{-2\pi i}{\lambda f} \tilde{r}\rho \cos(\tilde{\theta}-\phi)} E_{\text{out}}(\tilde{r}, \tilde{\theta}) d\tilde{\theta}\tilde{r} d\tilde{r} \\
&= \frac{1}{\lambda f} \sum_k \int \int e^{\frac{-2\pi i}{\lambda f} \tilde{r}\rho \cos(\tilde{\theta}-\phi)} e^{ikm\tilde{\theta}} E_{\text{out},k}(\tilde{r}) d\tilde{\theta}\tilde{r} d\tilde{r}.
\end{aligned} \tag{29}
$$

From the definition (Eq. (22)) of the Bessel functions as certain Fourier coefficients, we get

$$
\begin{aligned}
\int e^{ikm\tilde{\theta}} e^{\frac{-2\pi i}{\lambda f} \tilde{r}\rho \cos(\tilde{\theta}-\phi)} d\tilde{\theta} &= \int e^{ikm(\tilde{\theta}+\phi)} e^{\frac{-2\pi i}{\lambda f} \tilde{r}\rho \cos(\tilde{\theta})} d\tilde{\theta} \\
&= 2\pi i^{km} e^{ikm\phi} J_{km}\left(-\frac{2\pi}{\lambda f}\tilde{r}\rho\right).
\end{aligned} \tag{30}
$$

Substituting Eq. (30) into Eq. (29), we get

$$
\begin{aligned}
E_{\text{img}}(\rho, \phi) &= \frac{2\pi}{\lambda f} \sum_k e^{ikm\phi} \int i^{km} J_{km}\left(-\frac{2\pi}{\lambda f}\tilde{r}\rho\right) E_{\text{out},k}(\tilde{r})\tilde{r} d\tilde{r}, \\
&= \sum_k e^{ikm\phi} E_{\text{img},k}(\rho),
\end{aligned} \tag{31}
$$

where

$$
E_{\text{img},k}(\rho) := \frac{2\pi}{\lambda f} \int i^{km} J_{km}\left(-\frac{2\pi}{\lambda f}\tilde{r}\rho\right) E_{\text{out},k}(\tilde{r})\tilde{r} d\tilde{r}. \tag{32}
$$

Finally, since $E_{\text{img},k}(\rho) = E_{\text{img},-k}(\rho)$, we see that

$$
E_{\text{img}}(\rho, \phi) = E_{\text{img},0}(\rho) + 2\sum_{k=1}^{\infty} \cos(km\phi) E_{\text{img},k}(\rho). \tag{33}
$$

Note that

$$
|J_k(x)| \approx \frac{1}{k!}\left(\frac{x}{2}\right)^k
$$

for $0 \leq x \ll 1$. Hence, if we assume that $\epsilon < 10^{-3}$, then the error incurred by dropping all terms in the sum on k for which $k > 2$ will be approximately 10^{-6} at the most. Hence, any error in the PSF will be at the 10^{-12} level.

Thus, for a given Zernike polynomial aberration at the entrance pupil and ϵ in Eq. (21), Eq. (33) can be used to compute the resulting response of the apodized pupil mapping system. Note that an arbitrary aberration can be expressed as a linear combination of Zernike polynomials and, for each Zernike, the azimuthal variable can be integrated explicitly, thereby reducing 2D integrals to sums of 1D integrals.

3.2. Pure Apodization as a Special Case of Apodized Pupil Mapping

If we set $A_{\mathrm{pupmap}}(\tilde{r}) = 1$ for all $\tilde{r}$, then the pupil mapping system reduces to a trivial forward propagation of a flat beam. That is, the lenses/mirrors have perfectly flat surfaces and no remapping occurs. Hence, in this case, the hybrid system becomes a pure apodization system with the desired high-contrast apodization being achieved by a pair of apodizers, one appearing before the beam is propagated and one after (except for the Fresnel diffraction due to the propagation between the two mirrors.) In this fashion, the mathematics developed in 3.1 can be applied to a pure pupil-apodization coronagraph.

3.3. Off-Axis Diffraction Response

In addition to modeling optical aberrations, the equations in Section 3.1 can be used to model off-axis sources such as planets because they are mathematically equivalent to large tilt aberrations. For example, a 1au planet at 10pc is equivalent to a tilt aberration of 0.1 arcseconds. For a $D = 4$ meter aperture and $\lambda = 632.8$nm, this would correspond to an angle of $3\lambda/D$ and $\epsilon = 9.63$ in Eq. (21). Figure 3 is a contrast plot showing the result of computing Eq. (33) for such a planet, along with a simulation for a pure pupil-apodizing coronagraph. In addition, an on-axis star was added in both cases that is 10^8 times brighter than the planet. We assumed a coronagraph entrance aperture of 25mm, and a focal length of 2.5m, so that for the apodizing coronagraph, or any conventional imaging system, the planet appears at roughly 194μm off-axis in the image plane.

As mentioned before, pupil mapping both magnifies, which is good, and distorts, which is bad. Both effects are evident in Figure 3. The magnification moves the center of the distorted planet image out by roughly a factor of 2 (to about 400μm in Figure 3). Thus, the inner working angle of the system is effectively reduced, which is one of the benefits of pupil mapping. On the other hand, pupil mapping distorts the planet, which reduces the peak intensity a little with respect to the star's peak intensity. However, note that Figure 3 shows contrast plots, i.e. intensity is normalized to peak star intensity for each case. On an absolute intensity scale, the whole pupil mapping curve will be higher because of the throughput advantage. Also note that the distortion caused by the two-mirror system can be countered by inverting the pupil mapping after starlight has been blocked, as described at the end of Section 2.3.

3.4. Analysis of Sensitivity to Zernike Aberrations

Figures 4 to 7 show sensitivity results for the apodized pupil mapping system described in Section 2.3. Figure 4 is for aberrations that are 1/100-th wave rms whereas Figure 5 is for smaller 1/1000-th wave rms aberrations. Figure 6 shows radial cross-sections and Figure 7 shows differences between the radial cross-sections of the aberrated and un-aberrated PSF's. (As noted, the $(0,0)$, or piston, Zernike is differenced with the ideal PSF.) Note that these plots show the PSF for an on-axis source, i.e. a star. By comparing these plots to Figure 3, one can directly compare the levels of the aberrations to the PSF of an off-axis planet.

Two important effects are evident in Figure 7. First, the $(0,0)$, or piston, Zernike represents the nominally achieved contrast and so its difference with the ideal shows the limitation of the current design approach. Any additional errors due to the Zernike aberrations, as shown in the subsequent subplots (which are differenced with the nominal $(0,0)$ PSF), add to this static limit. Second, the pupil-mapping system shows a degradation in inner working angle with increasing Zernike and increasing aberration size. Any design that attempts to exploit the inherent magnification to achieve a smaller inner working angle must be able to achieve a higher level of aberration control.

Another important issue to consider is the sensitivity to aberrations at wavelengths other than the design wavelength. Comparing Figure 6 with the bottom-right subplot of Figure 2, it is evident that the degradation due to aberrations at 1/1000 wave level are generally higher than the degradation due to detuning of the wavelength by a factor of 1.3 or 0.7. Thus, the sensitivity to aberrations at these wavelengths should look very similar to Figure 6. Our simulations do indeed confirm this. Of course, at the longer wavelength, the aberrated PSF in Figure 6 would be expanded by the appropriate amount, and contracted for the shorter wavelength.

4. Sensitivity Analysis for Concentric Ring Masks

For comparison purposes, we now consider a high-contrast imaging system based on a simple concentric rings shaped pupil as described in Vanderbei et al. (2003b). A concentric-ring pupil-plane mask can be thought of as a circularly symmetric binary apodization. Hence, we assume that we are given such an apodization function $A(r)$.

As before, we assume that the input electric field could involve some aberration given by $\alpha(r,\theta)$. The image plane electric field is given by the Fourier transform of the pupil-plane

field, which in polar coordinates is

$$E_{\text{img}}(\rho, \phi) = \frac{1}{\lambda f} \int \int e^{-\frac{2\pi i}{\lambda f} r\rho \cos(\theta - \phi)} e^{i\alpha(r,\theta)} A(r) d\theta r dr. \tag{34}$$

Expanding the two exponentials in a Jacobi-Anger expansion, we get that

$$E_{\text{img}}(\rho, \phi) = \frac{1}{\lambda f} \int_0^a \sum_j \sum_k i^{j+k} J_j \left(-\frac{2\pi}{\lambda f} r\rho\right) J_k \left(\epsilon Z_l^m(r/a)\right) e^{-ij\phi} \left[\int_0^{2\pi} e^{ij\theta + km\theta} d\theta\right] A(r) r dr. \tag{35}$$

The integral in brackets is easily solved, as before, eliminating one of the summations. With some rearranging, this yields

$$E_{\text{img}}(\rho, \phi) = \frac{2\pi}{\lambda f} \sum_k i^{(1-m)k} e^{ikm\phi} \int_0^a J_{km} \left(\frac{2\pi}{\lambda f} r\rho\right) J_k \left(\epsilon Z_l^m(r/a)\right) A(r) r dr. \tag{36}$$

Finally, we use the symmetry property of the Bessel Functions $(J_{-m}(x) = (-1)^m J_m(x))$ to yield the image plane field,

$$E_{\text{img}}(\rho, \phi) = E_{\text{img},0}(\rho) + 2 \sum_{k=1}^{\infty} \cos(km\phi) E_{\text{img},k}(\rho), \tag{37}$$

where

$$E_{\text{img},k}(\rho) = \frac{2\pi}{\lambda f} \int_0^a i^{(1-m)k} J_{km} \left(\frac{2\pi}{\lambda f} r\rho\right) J_k \left(\epsilon Z_l^m(r/a)\right) A(r) r dr. \tag{38}$$

For a concentric ring mask having its edges at r_j, the formula for $E_{\text{img},k}(\rho)$ simplifies to

$$E_{\text{img},k}(\rho) = \frac{2\pi}{\lambda f} \sum_j \int_{r_{2j}}^{r_{2j+1}} i^{(1-m)k} J_{km} \left(\frac{2\pi}{\lambda f} r\rho\right) J_k \left(\epsilon Z_l^m(r/a)\right) r dr. \tag{39}$$

Figures 8 to 11 show sensitivity results for a high-contrast concentric-ring pupil mask. We also computed sensitivity results for the pure apodizing system; they appeared virtually identical to the results shown in these figures, which agree with the results in Green et al. (2004). Likewise, the off-axis results are indistinguishable from those of the pure apodizing system in Figure 3.

5. Conclusions

In this paper we presented an efficient method for calculating the distortions in the PSF of a pupil mapping system due to wavefront aberrations. Figures 4 to 11 show that our

particular pupil mapping system is somewhat more sensitive to low order aberrations than the concentric ring masks. That is, contrast and IWA degrade more rapidly with increasing rms level of the aberrations. However, this is partially mitigated by the magnification property of pupil mapping, making a direct comparison of the two systems more subtle, especially considering the distortion in pupil mapping. (A more direct comparison can be made for a pupil mapping that undoes the distortion, such as the unmapping system described at the end of Section 2.3, but that is outside the scope of this article.) In any case, it is evident that especially fine control of static and dynamic stability will be required in order to take full advantage of the smaller IWA intrinsic to the pupil mapping approach. A careful stability analysis of any design employing pupil mapping is thus necessary to determine its achievable operating range.

Finally, we note that there is a spectrum of apodized pupil mapping systems. The two extremes, pure apodization and pure pupil mapping, both have serious drawbacks. On the one end, pure apodization loses almost an order of magnitude in throughput and suffers from an unpleasantly large IWA. At the other extreme, pure pupil mapping fails to achieve the required high contrast. There are several points along this spectrum that are superior to the end points. We have focused on just one such point, which was suggested by Guyon et al. (2005). We leave it to future work to determine if this is the best design point. With this paper, we have provided the tools to analyze the sensitivity of these kinds of designs.

REFERENCES

M. Born and E. Wolf. *Principles of Optics*. Cambridge University Press, New York, NY, 7th edition, 1999.

P.S. Carney and G. Gbur. Optimal apodizations for finite apertures. *Journal of the Optical Society of America A*, 16(7):1638–1640, 1999.

A. Goncharov, M. Owner-Petersen, and D. Puryayev. Intrinsic apodization effect in a compact two-mirror system with a spherical primary mirror. *Opt. Eng.*, 41(12):3111, 2002.

J.J. Green, S.B. Shaklan, R.J. Vanderbei, and N.J. Kasdin. The sensitivity of shaped pupil

coronagraphs to optical aberrations. In *Proceedings of SPIE Conference on Astronomical Telescopes and Instrumentation*, 5487, pages 1358–1367, 2004.

O. Guyon. Phase-induced amplitude apodization of telescope pupils for extrasolar terrerstrial planet imaging. *Astronomy and Astrophysics*, 404:379–387, 2003.

O. Guyon, E.A. Pluzhnik, R. Galicher, R. Martinache, S.T. Ridgway, and R.A. Woodruff. Exoplanets imaging with a phase-induced amplitude apodization coronagraph–i. principle. *Astrophysical Journal*, 622:744, 2005.

J.A. Hoffnagle and C.M. Jefferson. Beam shaping with a plano-aspheric lens pair. *Opt. Eng.*, 42(11):3090–3099, 2003.

N.J. Kasdin, R.J. Vanderbei, D.N. Spergel, and M.G. Littman. Extrasolar Planet Finding via Optimal Apodized and Shaped Pupil Coronagraphs. *Astrophysical Journal*, 582: 1147–1161, 2003.

M.J. Kuchner, J. Crepp, and J. Ge. Finding terrestrial planets using eighth-order image masks. *Submitted to The Astrophysical Journal*, 2004. (astro-ph/0411077).

E.A. Pluzhnik, O. Guyon, S.T. Ridgway, R. Martinache, R.A. Woodruff, C. Blain, and R. Galicher. Exoplanets imaging with a phase-induced amplitude apodization coronagraph–iii. hybrid approach: Optical design and diffraction analysis. *Submitted to The Astrophysical Journal*, 2006. (astro-ph/0512421).

M. Shao, B.M. Levine, E. Serabyn, J.K. Wallace, and D.T. Liu. Visible nulling coronagraph. In *Proceedings of SPIE Conference on Astronomical Telescopes and Instrumentation*, number 61 in 5487, 2004.

D. Slepian. Analytic solution of two apodization problems. *Journal of the Optical Society of America*, 55(9):1110–1115, 1965.

W.A. Traub and R.J. Vanderbei. Two-Mirror Apodization for High-Contrast Imaging. *Astrophysical Journal*, 599:695–701, 2003.

R. J. Vanderbei, N. J. Kasdin, and D. N. Spergel. Checkerboard-mask coronagraphs for high-contrast imaging. *Astrophysical Journal*, 615(1):555, 2004.

R.J. Vanderbei. Diffraction analysis of 2-d pupil mapping for high-contrast imaging. *Astrophysical Journal*, 636:528, 2006.

R.J. Vanderbei, D.N. Spergel, and N.J. Kasdin. Circularly Symmetric Apodization via Starshaped Masks. *Astrophysical Journal*, 599:686–694, 2003a.

R.J. Vanderbei, D.N. Spergel, and N.J. Kasdin. Spiderweb Masks for High Contrast Imaging. *Astrophysical Journal*, 590:593–603, 2003b.

R.J. Vanderbei and W.A. Traub. Pupil Mapping in 2-D for High-Contrast Imaging. *Astrophysical Journal*, 626:1079–1090, 2005.
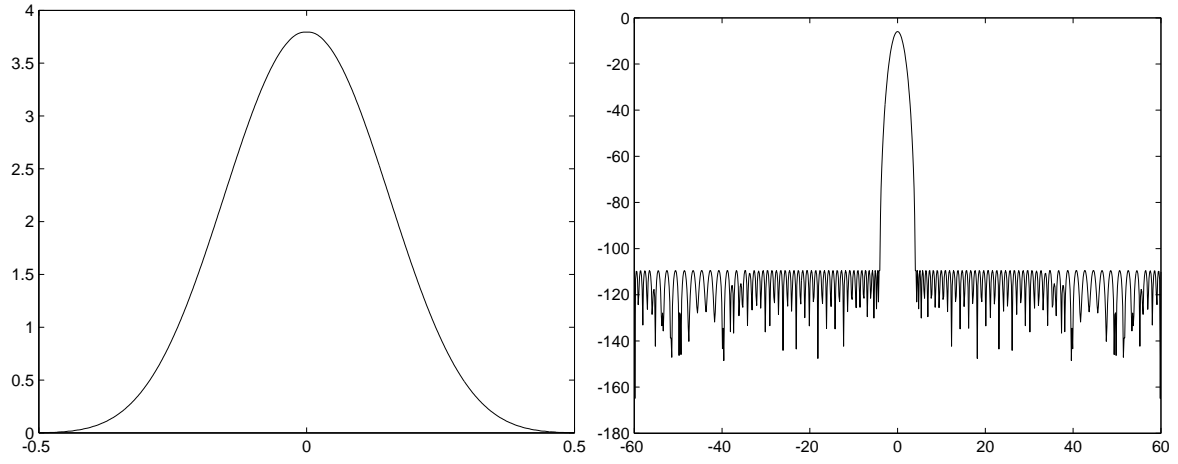
Fig. 1.— *Left.* An amplitude profile providing contrast of 10^{-10} at tight inner working angles. *Right.* The corresponding on-axis point spread function.
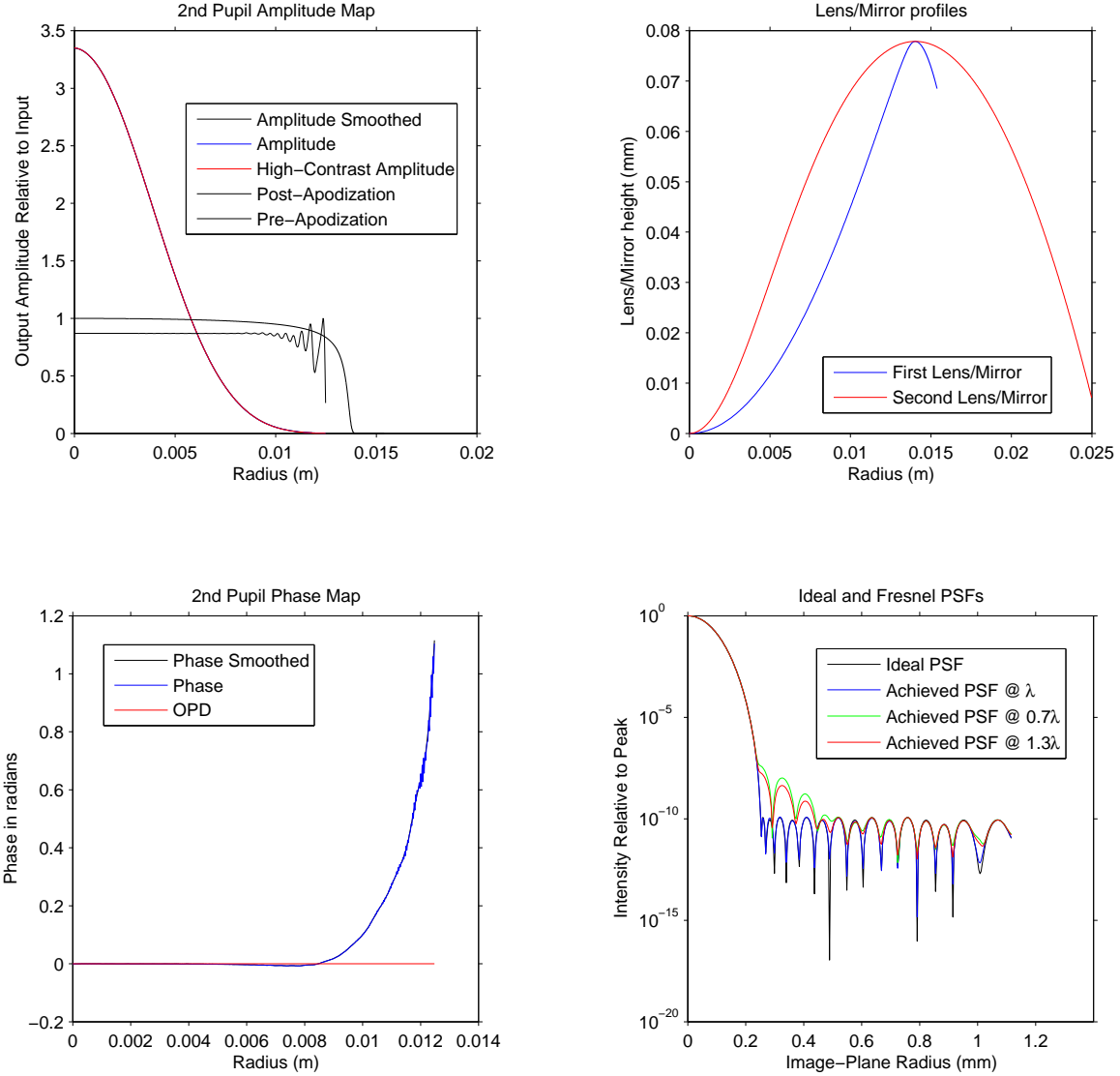
Fig. 2.— Hybrid pupil mapping system as described in Section 2.3. *Upper-left* plot shows in red the target high-contrast amplitude profile and in blue the Gaussian profile that is a close approximation to it. The two black plots show the pre- and post-apodizers. *Upper-right* plot shows the lens profiles, blue for the first lens and red for the second. *Lower-left* plot shows in blue the phase map computed using Huygens propagation and in black the smoothed approximation. *Lower-right* plot shows in black the ideal PSF, in blue the actual PSF at the design wavelength, in red the actual PSF at 130% of the design wavelength, and in green the actual PSF at 70% of the design wavelength. (The horizontal scale for 130% and 70% has been dilated by factors of $1/1.3$ and $1/0.7$, respectively, and the x-axis labels correspond to the original λ.)
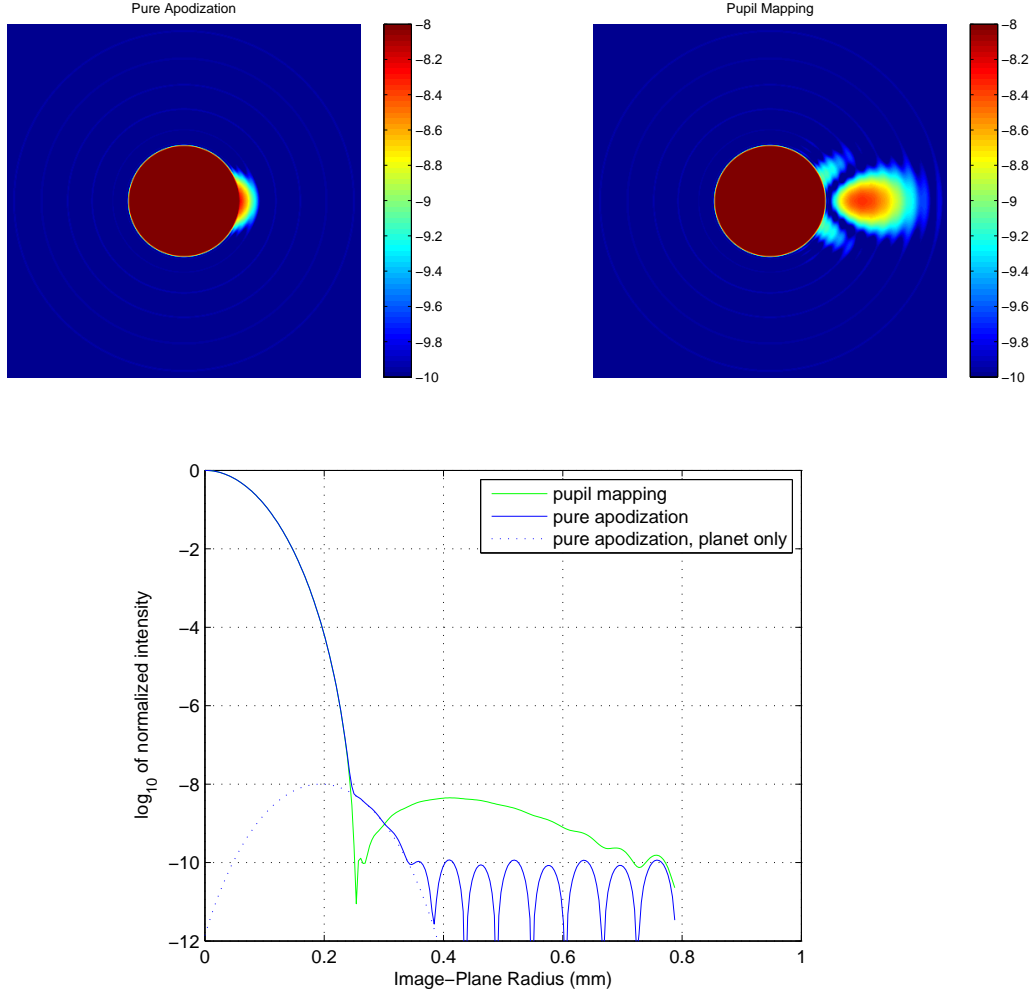
Fig. 3.— Distortion of off-axis sources in pupil mapping. *Top*: images of a $3\lambda/D$ planet (e.g. a 1au planet viewed from 10 parsecs away with a 4m telescope). The star appears as a saturated disk. Both images are contrast plots normalized to the peak intensity of the star image. The z scale is log_{10} of contrast. *Bottom*: Horizontal slices. Note that these are contrast plots, i.e. everything is normalized to the peak intensity of the star image in pupil mapping as well as pure apodization. On an absolute intensity scale, the whole pupil mapping curve would be higher due to better throughput.
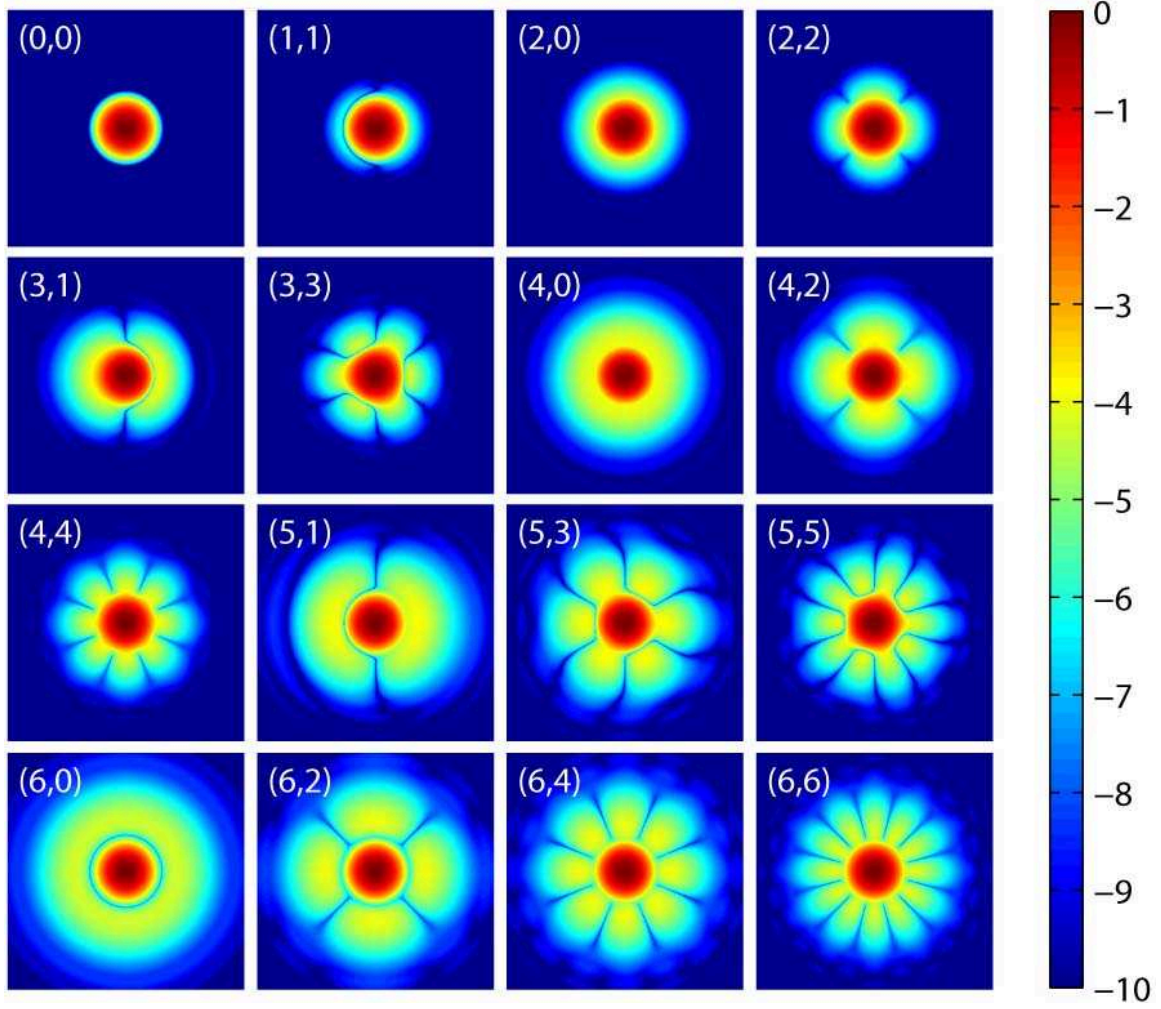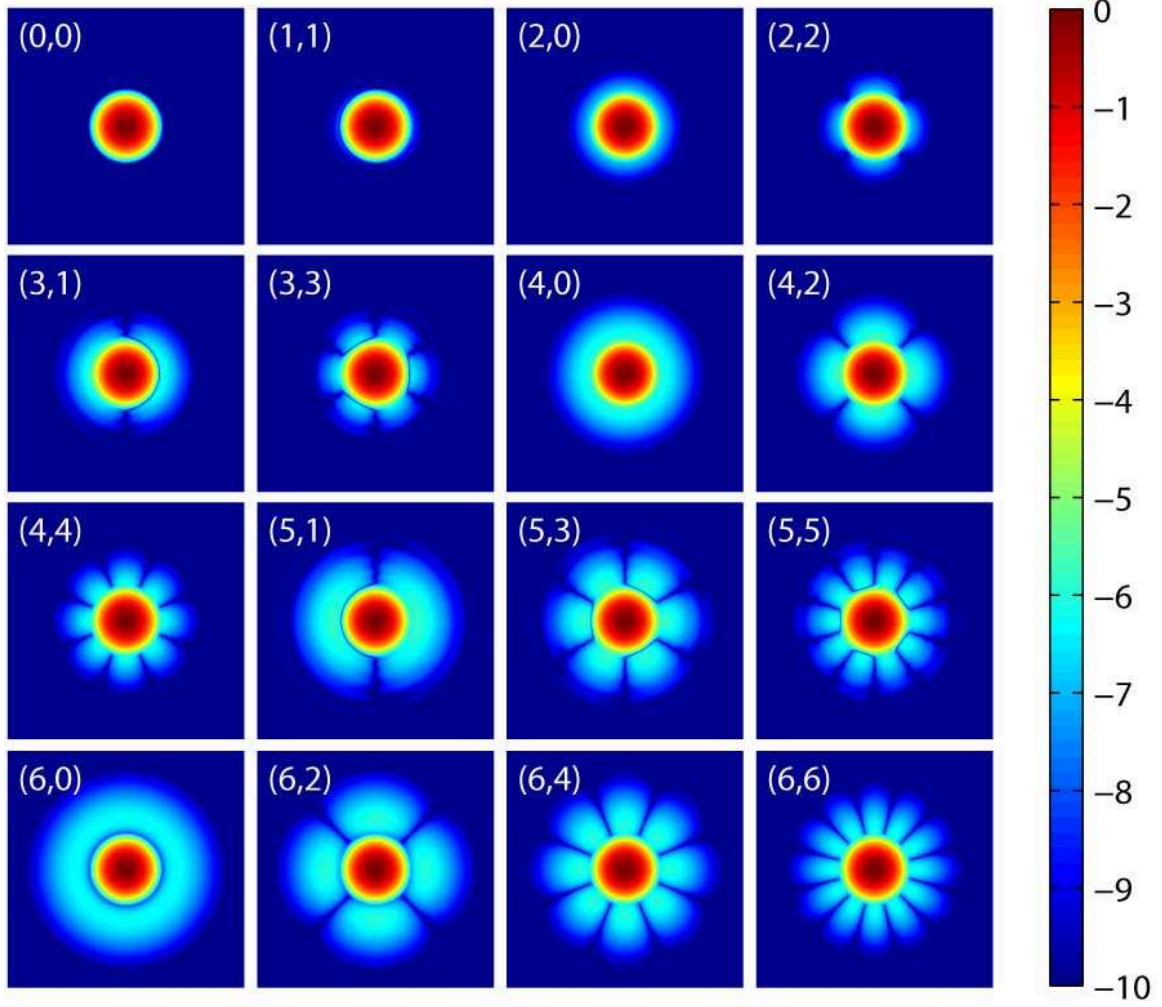
Fig. 4.— Pupil mapping. PSFs associated with various Zernikes normalized so that the rms error is, in each case, 1/100-th wave. The plots correspond to: Piston $(0,0)$, Tilt $(1,1)$, Defocus $(2,0)$, Astigmatism $(2,2)$ Coma $(3,1)$, Trefoil $(3,3)$, Spherical Aberration $(4,0)$, Astigmatism 2nd Order $(4,2)$, Tetrafoil $(4,4)$, etc. *Important note:* As explained in Section 2.3, off-axis sources are "pushed out" by a magnification factor of about 2 in these high-contrast pupil-mapping systems. This needs to be remembered when comparing with later plots.

Fig. 5.— Pupil mapping. PSFs associated with various Zernikes normalized so that the rms error is, in each case, 1/1000-th wave. The plots correspond to: Piston $(0,0)$, Tilt $(1,1)$, Defocus $(2,0)$, Astigmatism $(2,2)$, Coma $(3,1)$, Trefoil $(3,3)$, Spherical Aberration $(4,0)$, Astigmatism 2nd Order $(4,2)$, Tetrafoil $(4,4)$, etc.
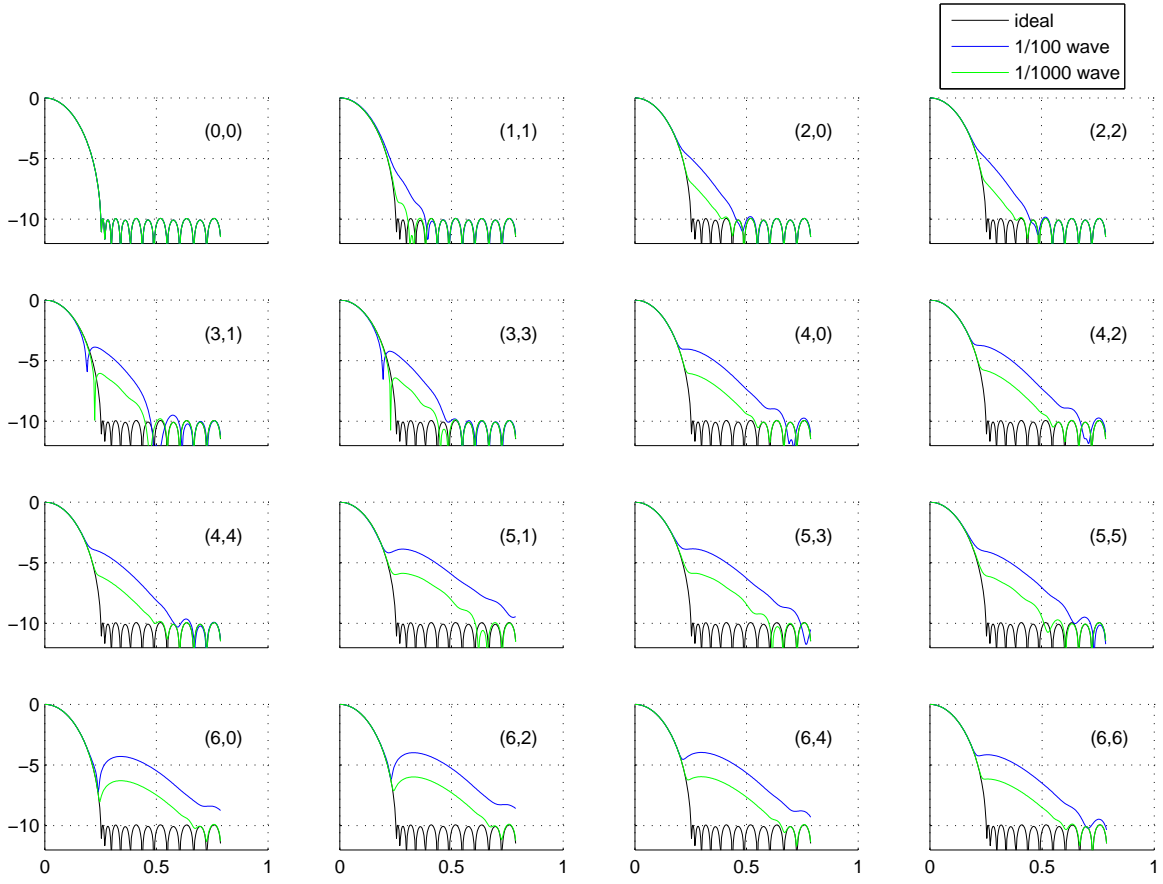
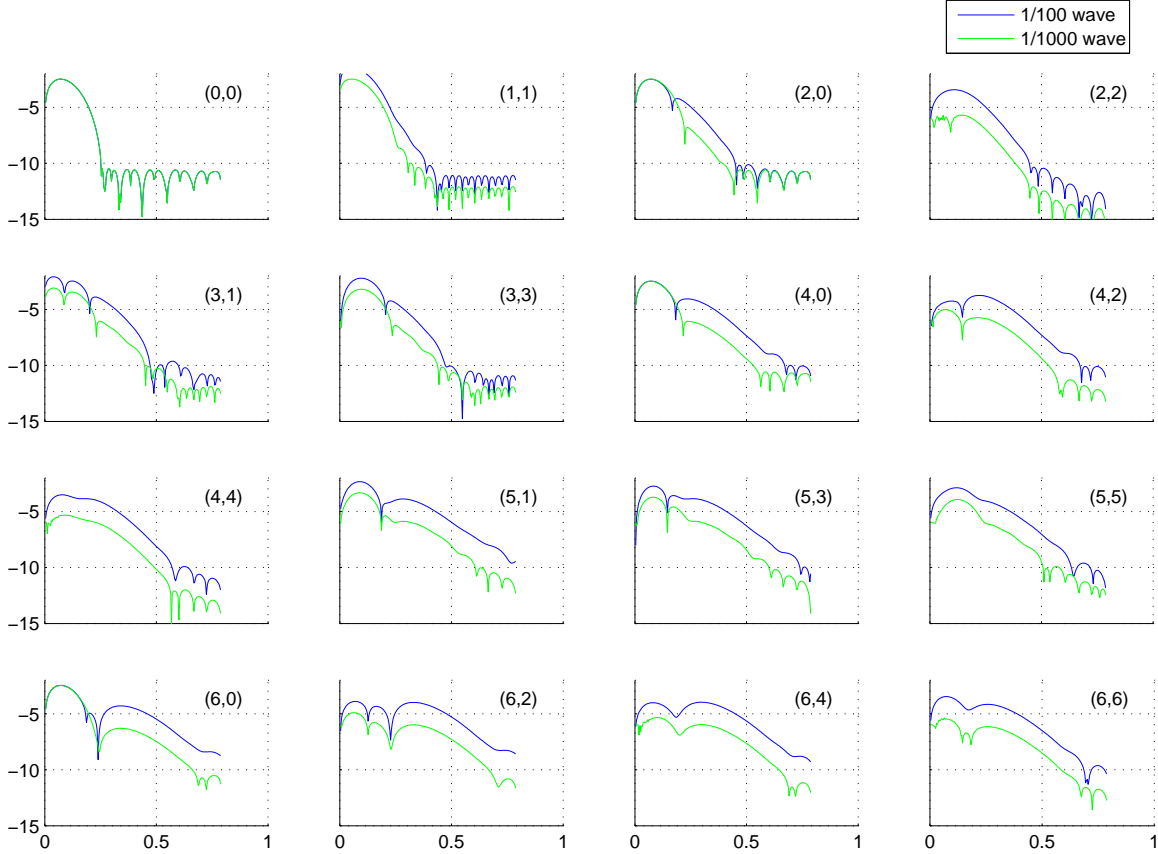Fig. 6.— Cross-sectional 1-D plots through the positive x-axis in the plots of the previous two figures.

Fig. 7.— Difference plots. For all the plots except (0,0), the difference is between the aberrated and non-aberrated profile for pupil mapping. (The non-aberrated profile is essentially the (0,0), or piston term). For the (0,0) difference plot in the upper left, the difference is between the non-aberrated pupil mapping profile and the non-aberrated ideal apodization profile, showing that pupil mapping does not match the ideal apodization profile at the contrast level of 10^{-11} and below.
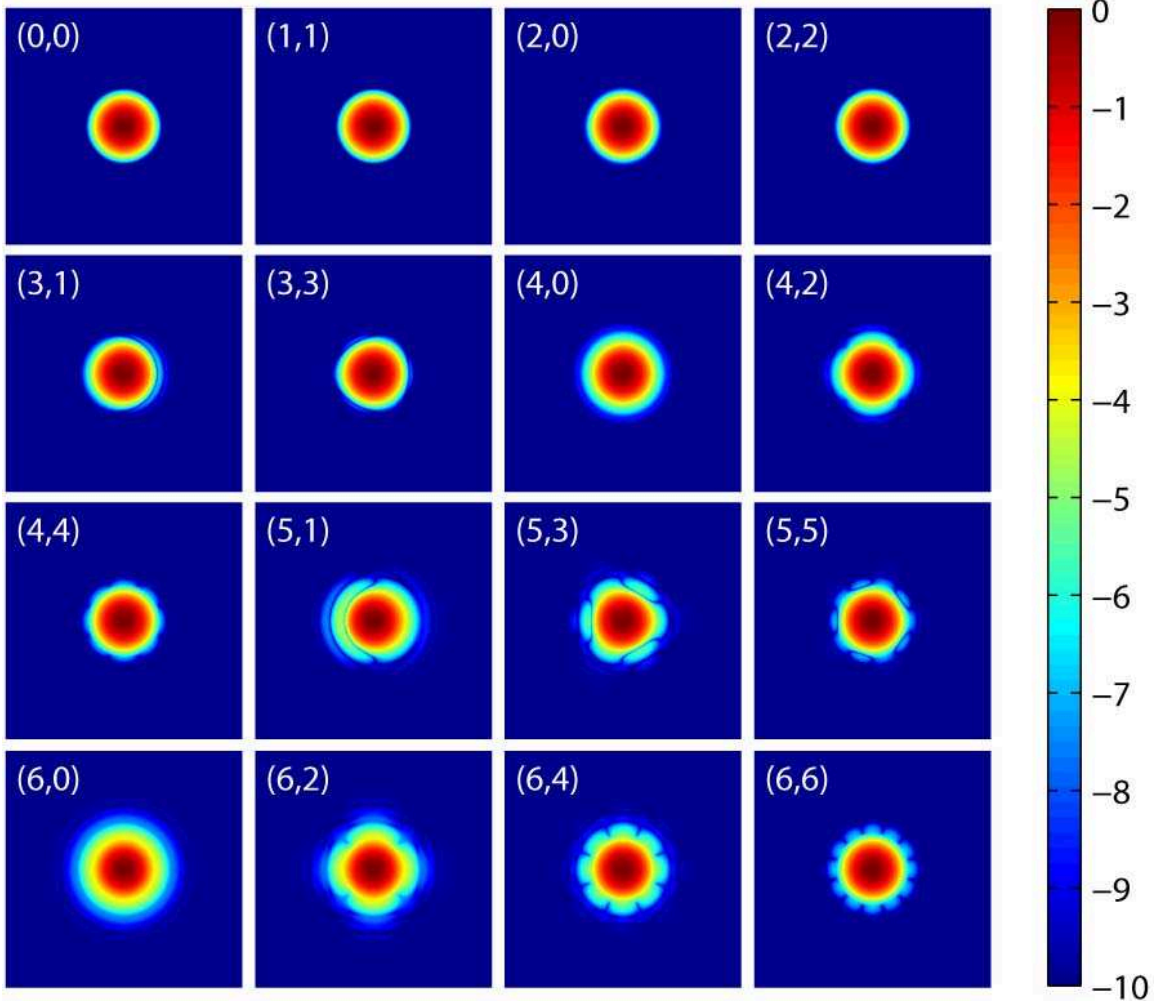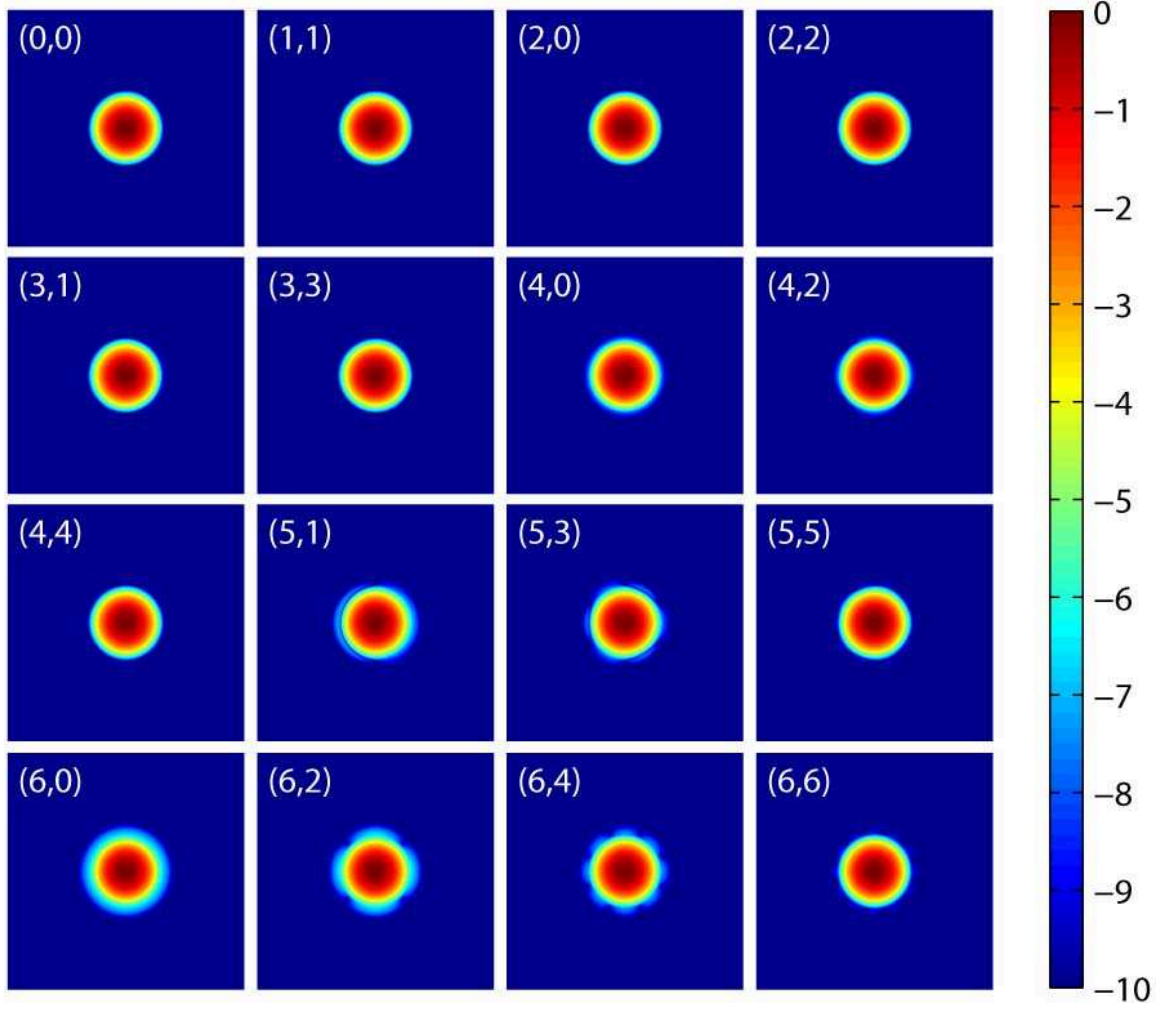
Fig. 8.— Concentric Rings. PSFs associated with various Zernikes normalized so that the rms error is, in each case, 1/100-th wave. The plots correspond to: Piston $(0,0)$, Tilt $(1,1)$, Defocus $(2,0)$, Astigmatism $(2,2)$ Coma $(3,1)$, Trefoil $(3,3)$, Spherical Aberration $(4,0)$, Astigmatism 2nd Order $(4,2)$, Tetrafoil $(4,4)$, etc.

Fig. 9.— Concentric Rings. PSFs associated with various Zernikes normalized so that the rms error is, in each case, 1/1000-th wave. The plots correspond to: Piston $(0, 0)$, Tilt $(1, 1)$, Defocus $(2, 0)$, Astigmatism $(2, 2)$ Coma $(3, 1)$, Trefoil $(3, 3)$, Spherical Aberration $(4, 0)$, Astigmatism 2nd Order $(4, 2)$, Tetrafoil $(4, 4)$, etc.
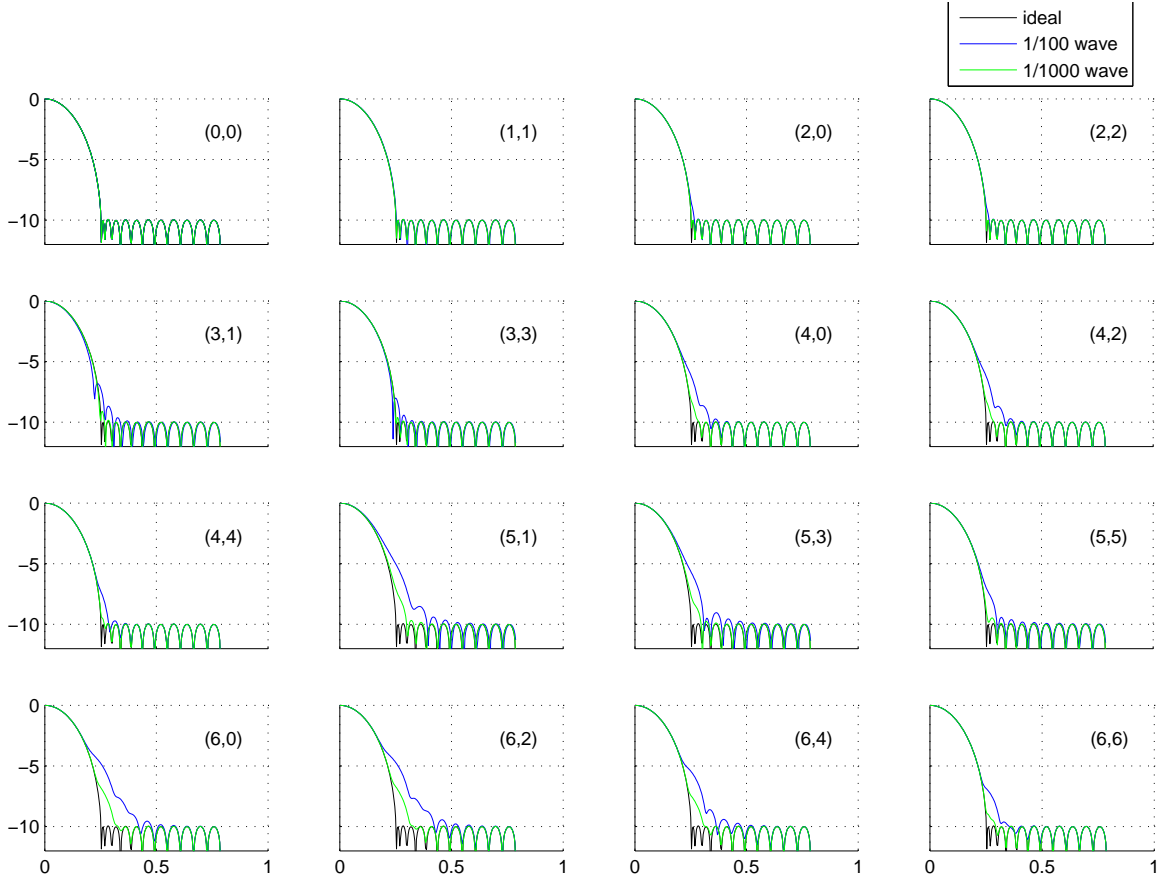
Fig. 10.— Cross-sectional 1-D plots through the positive x-axis in the plots of the previous two figures.
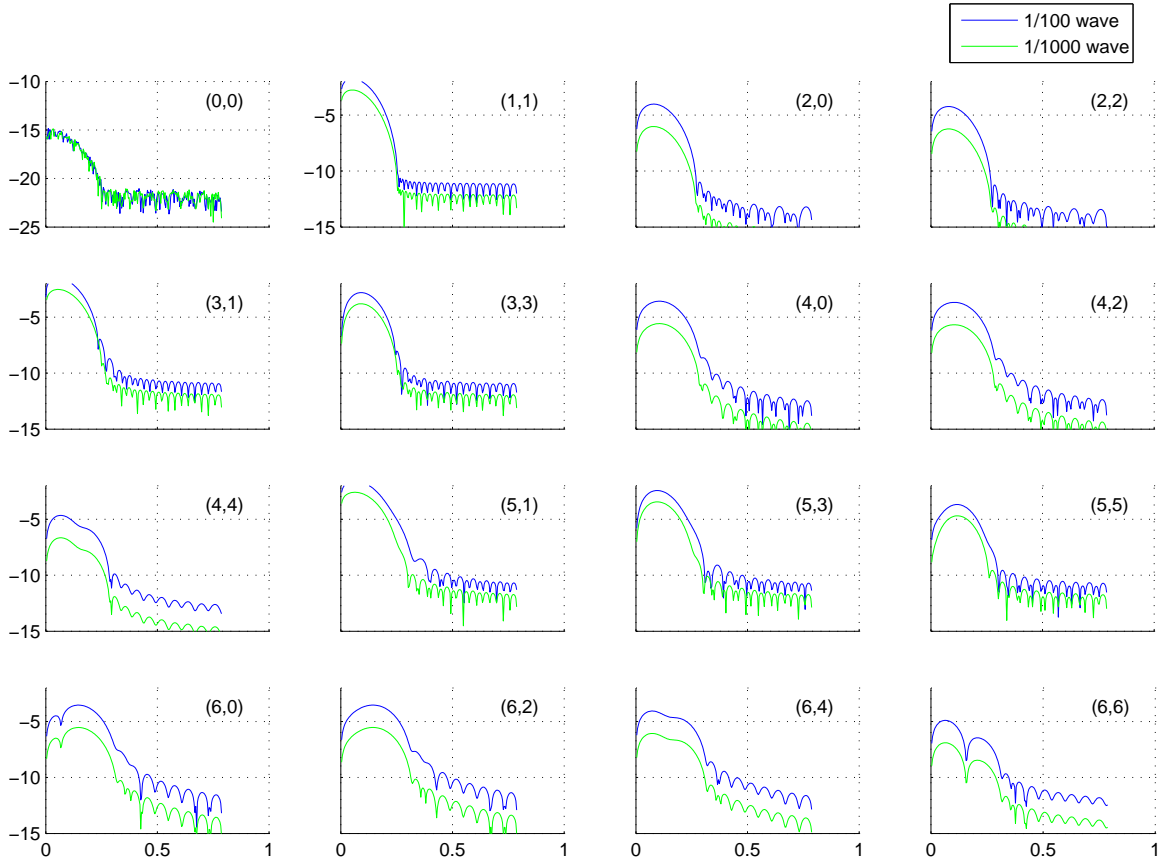
Fig. 11.— Difference plots. Differences between the aberrated profile and the corresponding ideal.